



Summary

Contents

A Quality Information Bulletin from the Methodology Division

Contents

Measures of Australia's Progress Takes Centre Stage at OECD World Forum
Improving Coherence Between Seasonally Adjusted Estimates
Guidelines for Clearance and Release of ABS Analytical Work
Sample Design for Movement Estimates in ABS Business Surveys
Census Imputation for Non-Responding Dwellings
Evaluating Labour Force Survey Non-Response Impact through Imputation
Sample Design for the Business Longitudinal Database
Unsentenced Prisoners in Australia

Measures of Australia's Progress Takes Centre Stage at OECD World Forum

MEASURES OF AUSTRALIA'S PROGRESS TAKES CENTRE STAGE AT OECD WORLD FORUM

Australian Bureau of Statistics (ABS) staff recently went to Palermo, Sicily to attend an OECD held world forum on Key Indicators: Statistics, Knowledge and Policy.

The Forum was attended by some 500 high level experts from OECD countries and a few non-OECD countries. They represented various social, economic, statistical, media, and political communities, as well as key sectors and disciplines and the media. The participants shared information and compared strategies for measuring and assessing the overall position and progress of a certain political entity (country, region, etc.) **vis-à-vis** other similar entities. The forum's purpose was to promote research and information sharing among those who work to promote the use and development of indicator systems.

As the OECD explained "in today's rapidly changing, increasingly interdependent world, productive debate and decisions require information that is comprehensive, trustworthy and easy to understand. And while many countries are developing and maintaining both specialised and comprehensive indicator systems, there has not been any co-ordinated world-wide effort to study how to develop and use these large-scale systems of public information."

The ABS Measures of Australia's Progress (MAP) project was used as one of four keynote case studies at the forum, and was held up as a successful model that other countries might aspire to. The Australian Statistician, Dennis Trewin, gave a plenary address on the project, while Jon Hall from ABS Methodology Division presented a paper on civil society's role in developing indicator projects (including MAP) in Australia.

For further information, please contact, Jon Hall on (02) 6252 7221

Email: jon.hall@abs.gov.au

Improving Coherence Between Seasonally Adjusted Estimates

IMPROVING COHERENCE BETWEEN SEASONALLY ADJUSTED ESTIMATES

The Australian Bureau of Statistics (ABS) publishes seasonally adjusted indicators for many subject matter area (SMA) collections. A lack of coherence in the seasonal adjustment of broadly similar original estimates can be

observed when related collections are independently seasonally adjusted, resulting in two different sets of seasonally adjusted movements, which on some occasions are in opposite directions.

It would generally be expected that broadly similar original estimates from different ABS SMA's (inputs) would result in broadly similar seasonally adjusted indicators (outputs). However, discrepancies between related collections' original estimates and between seasonally adjusted methods have resulted in different seasonally adjusted movements. Identifying and addressing the causes of these discrepancies should improve coherence of these seasonally adjusted movements.

The independent seasonal adjustment of similar ABS collections was found to be limited to a total of twenty cases. Nineteen of these cases related to National Accounts time series and their corresponding time series from other SMA's, especially Quarterly Business Indicators Surveys and Construction Surveys.

Of these nineteen cases, the major cause was found to be discrepancies between the original estimates, which, although based on ABS survey output, were sometimes subject to further adjustments due to appropriate data confrontation or to transforming the source SMA's estimates to a "National Accounts" basis.

All twenty cases also had some discrepancy between their seasonal adjustment methods. This was due to differences in the seasonal adjustment approach (concurrent or forward factor), type (direct or indirect) or prior corrections applied, especially trading day, Easter, structural breaks and outlier corrections.

In conclusion, it was recognised that discrepancies in the methods of seasonal adjustment between related collections should be addressed to improve coherence. Time Series Analysis section has begun to work closely with National Accounts and other SMA's to improve coherence in recent annual seasonal reanalyses. In the long term, a more systematic approach will be pursued, by linking meta data between related time series to improve their seasonal adjustment coherence.

For further information, please contact, Tom Outteridge on (02) 6252 6406

Email: tom.outteridge@abs.gov.au.

Guidelines for Clearance and Release of ABS Analytical Work

GUIDELINES FOR CLEARANCE AND RELEASE OF ABS ANALYTICAL WORK

Putting analytical work and findings into the public domain and making it accessible via the ABS website has been an ongoing challenge for Analytical Services Branch. Guidelines for the clearance and release of analytical work were recently prepared by the Branch, discussed within Methodology Division (MD), and presented to the Analysis Board.

The guidelines build on work undertaken within Information Management Division on classifying ABS products. Most analytical work will fall within the ABS Research Paper category. Three checklists have been developed for reviewing and ultimately clearing analytical research papers prior to making them available in the public domain. These are:

- a checklist for review within the branch and possibly the wider Division,
- a checklist for review by internal clients outside of the Division, and
- a checklist for review by external clients such as for methodological rigour.

Comments and suggestions from within MD and from the Analysis Board have been taken on board to revise the set of guidelines. We are now planning on trialling the guidelines on a number of products across MD, such as papers for the forthcoming 2005 International Statistical Institute (ISI), Conference for external conferences, and for Methodology Advisory Committee meetings.

In addition, MD is also looking at the location of its work in the ABS website, the format of its research papers, and the classification of analytical products on the website. As a package, we are expecting to make substantial inroads into making more analytical work visible.

For further information, please contact, Marion McEwin on (02) 6252 7290

Email: marion.mcewin@abs.gov.au.

Sample Design for Movement Estimates in ABS Business Surveys

SAMPLE DESIGN FOR MOVEMENT ESTIMATES IN ABS BUSINESS SURVEYS

Although most Australian Bureau of Statistics (ABS) repeated business surveys have historically been designed to produce reliable point-in-time estimates, the users of these surveys are often more interested in the movement estimates (for example, the principal objective of the Retail Trade Survey is to show month to month movement of turnover for retail and selected service industry). The reliability of the movement estimates have primarily been controlled through the selection of the samples, using a method of permanent random number sampling (referred to as synchronised sampling) control overlap of samples between periods. Typically a certain percentage of units will be rotated out of the sample each period since there is a requirement to minimise the provider load on individual businesses over time. The reliability of the movement estimates have also been reliant on the assumption that the optimal sample allocations for the movement estimates will be similar to the optimal sample allocations for the point-in-time estimates.

An evaluation was undertaken to determine whether the reliability of the movement estimates can be improved through allocation sample between strata. The evaluation considered three alternative sample allocations:

- optimal Point-in-Time Sample Allocation - Equal State relative standard errors on the point-in-time estimates,
- optimal Movement Sample Allocation - Equal State relative standard errors on the movement estimates,
- optimal Compromise Sample Allocation - Equal State relative standard errors on the movement estimates and (fifty percent larger) equal State relative standard errors on the point-in-time estimates.

The evaluation using the Survey of Average Weekly Earnings found that:

- the differences between the stratum sample sizes under the various sample allocations were quite substantial,
- the **optimal point-in-time** and **optimal compromise** sample allocations produced more reliable point-in-time estimates than the **optimal movement** sample allocation.
- the **optimal movement** and **optimal compromise** sample allocation failed to produced more reliable movement estimates than the **optimal point-in-time** sample allocation, primarily because the movement stratum population variances were more variable across the reference periods than the point-in-time stratum population variances.

The results of the evaluation indicated that the reliability of the movement estimates cannot be improved by designing for movement estimates, and that the current method of designing to produce reliable point-in-time estimates is the best approach for designing ABS repeated business surveys.

For more information, please contact, John Preston on (02) 6252 6970.

Email: john.preston@abs.gov.au.

Census Imputation for Non-Responding Dwellings

CENSUS IMPUTATION FOR NON-RESPONDING DWELLINGS

A review of the 2001 Census imputation methodology was conducted. In particular the review focussed on removing the processing induced overcount of roughly 83 000 persons.

When a Census collector believes a dwelling to be occupied, but (for a variety of reasons) no form is obtained for it then the dwelling is considered to be non-responding. Census records are created for these dwellings and for the persons believed to be living in them. These are called 'system created records' (previously they were known as 'dummy' records').

For most non-responding dwellings the number of males and females resident on Census night is imputed. In 2001 (and previously) this was done using the average number of males and females per household for the CD in which the dwelling was located. The average at the Australia level was 2.4 persons (both male and female).

Some non-responding dwellings supply their own count of number of males and females. This is acquired by asking a neighbour or by some other method, and is referred to as 'credible source data'. Analysis of credible source data from 2001 showed that these records differed from the general responding Census population. In particular, the average number of persons per household about which credible source data was available was 1.8. Given that there are roughly 140 000 system created dwellings this difference (1.8 versus 2.4) produces the estimated overcount.

The outcome of the review was that a new imputation methodology would be devised for non-responding dwellings that incorporated the credible source data as the best representation of non-respondents available. Following on

from this analysis a variety of imputation methods were assessed, including:

- using credible source data to create an adjustment factor for the CD level mean number of persons per responding household,
- using a model incorporating dwelling and CD characteristics to predict the number of persons per household,
- using the mean number of persons per household from credible source data within broad imputation classes,
- using a hotdecking methodology to select a donor with credible source data that had the same dwelling and geographic characteristics as a non-respondent.

The assessment involved reimputing 2001 non-respondents. The results were compared with the actual imputed values produced during 2001 Census processing.

All of the methods reduced the overcount by a desirable amount, and had similar levels of accuracy. The hotdecking method was ultimately selected because it imputes non-respondents with the same distribution of dwelling size as the credible source data. This is desirable because credible source records have a far higher proportion of one person households than the general population. Using hotdecking we impute 65 000 more one person households than were actually imputed for the 2001 Census. This has flow-on effects when we go on to impute other characteristics for system created records, such as age and marital status (ie small households are more likely to contain persons aged 20-35 and 55+ and less likely to contain persons aged 0-15).

Currently Population Census are working on a system to implement this method in readiness for testing in the 2005 Census Dress Rehearsal.

For further information, please contact, Claire Clarke on (02) 6252 5556

Email: claire.clarke@abs.gov.au

Evaluating Labour Force Survey Non-Response Impact through Imputation

EVALUATING LABOUR FORCE SURVEY NON-RESPONSE IMPACT THROUGH IMPUTATION

Operational difficulties during enumeration of the August 2004 Labour Force Survey (LFS) resulted in a much higher number of non-contact households than normal. The August national response rate was just over 94%, more than 2% below the national response rate typically achieved for the LFS.

Analyses were conducted to investigate whether the increased non-response unduly impacted on key estimates. Estimates would have been affected if, on average, the additional non-respondents in August had different labour force characteristics to the persons in the responding sample. Any such differences would have been realised as a change to the non-response bias typically experienced by the LFS, and create a spurious effect in estimates.

A longitudinal imputation procedure was used to estimate the change in non-response bias for labour force status estimates. Since the LFS has high overlap between the set of dwellings selected in consecutive months, it is possible to identify many individuals who were likely in coverage but did not respond to the survey. The imputation procedure imputed persons in non-contact and refusal households who responded to the survey in the previous month. The imputed labour force status was the individual's labour force status in the previous month. Analysis showed the strategy of carrying over the labour force status from the previous month is fairly reliable, with highest bias occurring in December and January when many persons enter and leave the labour market. Non-response bias was estimated as the difference between the original estimate and estimate obtained when the imputed persons are included for estimation.

The imputation procedure was applied across the previous 12 months to investigate the recent behaviour of the non-response bias estimates for employed and unemployed. The time series of the non-response bias for the Australian estimate of employed persons was relatively stable until August, when there was a statistically significant change in the bias. The change in bias indicated the August seasonally-adjusted estimate of employed persons was understated by between 15,000 and 25,000 persons. The non-response bias was reflected in the employed estimates at the State and Territory level, with the effect most pronounced in NSW, where August response rates were particularly low. The additional non-response did not significantly affect estimates of unemployed at the Australian or State levels. The imputation results imply that compared to the respondents, the additional non-respondents in August were more likely to be employed and less likely to be not in the labour force.

In the past two months response rates have approached usual levels, and accordingly non-response bias estimates have been close to normal. However, the change back to regular non-response bias levels in September impacted on August-September movements. The imputation procedure could be used in the future on an on-going basis for quality monitoring of the LFS.

For further information, please contact, Julian Whiting on (08) 8237 7362

Email: julian.whiting@abs.gov.au.

Sample Design for the Business Longitudinal Database

SAMPLE DESIGN FOR THE BUSINESS LONGITUDINAL DATABASE

As a response to strong external demand for longitudinal statistics on business activity and performance, the Australian Bureau of Statistics (ABS) has undertaken to create a Business Longitudinal Database (BLD). One of the primary aims of the BLD was to use data from existing sources (both internal and external such as Australian Taxation Office (ATO) data), however after consultation with potential users the need for data relating to business characteristics was emphasised. This meant that to populate the BLD with relevant data, a specific collection will be required to supplement data available from current ABS characteristic surveys, such as the new Innovation survey.

Based on discussions with external and ABS analysts, the preference for the BLD is a smaller-denser dataset rather than a larger-sparser dataset. That is, a dataset which is smaller in terms of sample size yet denser in terms of data items. This is because the richness of data across time from particular businesses is seen as more desirable than large samples with less data. In such a dataset all data items are available for all selected businesses for all time periods covered by the survey (the 'want it all principle'). This is a rather broad-brush description however, which masks the dilemma for dataset designers who need to resolve a number of difficult issues. Even if it is accepted that the ideal longitudinal dataset is one which is dense with information, there are many other characteristics that describe the dataset that also need to be considered. These include the sample size, sample distribution, number, spread and complexity of data items, number of waves, reference period covered and interval between waves.

Many of these will have no one solution, and in different situations a decisions on what is needed to produce a "good" dataset will vary. Decisions were made for the BLD based on discussion with various analysts, and with consideration of practical influences such as ABS provider load policy.

The final design determined for the BLD is an equal allocation at the Australia, New Zealand Standard Industry Classification (ANZSIC) Division by broad size level. The initial sample taken in the first year will be large enough to allow for deaths over time, so that by the end of the fifth year there will be sufficient live units available for analysis. The size of this sample will be determined by further consultation and budget constraints. The proposal to retain units in sample for five years was made with consideration to the minimum useful time suggested by analysts, and the current ABS policy on provider load. In each subsequent year a new panel of units will be added to the BLD and then followed for five years.

This method of adding a new panel each year ensures that for any five year period of interest there will be units available for analysis, but also relieves provider burden by not retaining the same units in the BLD year after year. The aim is as much as possible that the units included in the BLD will be taken from ABS characteristic surveys to minimise the size of the specific BLD collection.

For further information, please contact, Helen Teasdale on (08) 9360 5991

Email: helen.teasdale@abs.gov.au.

Unsentenced Prisoners in Australia

UNSENTENCED PRISONERS IN AUSTRALIA

The increasing unsentenced prisoner population has been of great interest in recent years, especially to policy makers, researchers and the wider community. The ratio of unsentenced prisoners to total prisoners has steadily increased over the last ten years, from 12% in 1993 to 20% in 2003. The Analysis Branch and National Centre for Crime and Justice Statistics is currently undertaking a project to investigate characteristics of the unsentenced prisoner population and the relationship between the size of the unsentenced prisoner population across states and time. Unsentenced prisoners include those who are unconvicted and awaiting a court hearing, convicted prisoners awaiting sentencing, and persons awaiting deportation.

Administrative data from **Corrective Services, Australia** (ABS cat. no. 4512.0) and **Prisoners in Australia**, (ABS cat. no. 4517.0) is being used to conduct both descriptive and multivariate analysis. A Poisson regression model has been used to better understand factors associated with the unsentenced prisoner population count assuming it has a Poisson probability distribution. The regression analysis showed that many operational, legislative and

demographic differences between states and over time are significant in explaining changes in the unsentenced prisoner population.

The preliminary results from this analysis were recently presented at an Australian Institute of Criminology Conference. One of the issues raised at the conference about the analysis was related to the difficulties of analysing administrative data, especially data which is censored (the Prisoner Census gathers information on prisoners who are in custody on the 30th June each year, so prisoners who are not there on that day will be excluded from the data set). These difficulties had already been addressed in the conference paper as well as our presentation.

We are currently conducting further research and analysis to gain a deeper understanding of the unsentenced prisoner population and aim to present the final results in an ABS Research Paper which will be made available from the ABS website.

For further information, please contact, Sarah Dexter on (02) 6252 7246

Email: sarah.dexter@abs.gov.au.

About this release

Methodological News

© Commonwealth of Australia

All data and other material produced by the Australian Bureau of Statistics (ABS) constitutes Commonwealth copyright administered by the ABS. The ABS reserves the right to set out the terms and conditions for the use of such material. Unless otherwise noted, all material on this website - except the ABS logo, the Commonwealth Coat of Arms, and any material protected by a trade mark - is licensed under a Creative Commons Attribution 2.5 Australia licence